

Student name: Emile Isopescu

Student number: 2004945

BA Course Title: Film & TV

Tutor name: Rob Jones

Word Count: 1647

Task: Observational Writing

In this essay, I will be writing about the philosophical interpretation of the 'self' through the lens of artificial intelligence, specifically how Philip K. Dick's novel *Do Androids Dream of Electric Sheep?* (Dick, 1968) explores this problem. The story of Deckard and his job of finding and destroying rogue androids (or 'andys'), artificially intelligent machines that resemble humans, asks a plethora of questions about what it means to be 'human,' and indeed the boundaries that need to be surpassed in order to achieve the manufacture of human intelligence outside of the organic brain.

We live in an age of rapid technological advancement, especially in the field of artificial intelligence. Our lives are now, more than ever, governed by algorithms designed to predict our interests across all walks of life; music, books, TV shows, even products that you might buy are suggested to us by code that runs automatically, independent of the people that initially developed and wrote it. It is my personal belief that it's only a matter of time before these systems and programs become indistinguishable from that of the organic processes that govern the functions of the human brain; in other words, we will eventually be creating technology that thinks and operates as human beings do. The ultimate question, then, is how far can we go before we consider these machines as human in and of themselves?

Do Androids Dream of Electric Sheep? Opens with the introduction of the mood organ, a device that artificially regulates human emotion. Protagonist Rick Deckard is debating with his wife about what moods they should 'dial in' for the day.

"He sighed, defeated by her threat. "I'll dial what's on my schedule for today."

Examining the schedule for January 3, 1992, he saw that a business-like professional attitude was called for. "If I dial by schedule," he said warily, "will you agree to also?" He waited, canny enough not to commit himself until his wife had agreed to follow suit.

"My schedule for today lists a six-hour self-accusatory depression," Iran said.

"What? Why did you schedule that?" It defeated the whole purpose of the mood organ." (Dick, 1968, p.2)

Immediately, Dick has begun to paint a bleak picture of the state of humanity in this world. The experience of feeling emotions naturally in response to external stimuli has been stripped away in favour of a schedule, an automated process governed by a machine that essentially feels things for you. The mood organ was ostensibly created to improve emotions, however this exchange between Rick Deckard and his wife Iran makes it clear that some people have reached a point where they use it to deliberately induce negative emotional states as well. This establishes a hugely important theme that runs throughout the entire narrative: in a time where highly advanced artificial intelligence exists, humans, too, have become somewhat machine-like.

The novel continues to heavily blur the line between human and machine with the androids; these highly advanced robots are physically identical to humans and are able to think and function just as we do to the extent that it is nearly impossible to distinguish them from humans. Initially built to serve us as slaves, androids eventually became so intelligent that they started to defect and rebel from their human masters, often in violent ways. Bounty hunters such as Rick Deckard are hired to identify and destroy wanted androids, and they're largely treated as subhuman. As Rick says in the book, "a humanoid robot is like any other machine; it can fluctuate between being a benefit and a hazard very rapidly. As a benefit, it's not our problem." (P. 34)

Their subhuman status is further highlighted by the fact that destroying an android is not referred to as 'killing', but rather 'retiring,' a term that has clearly been used to remove the moral weight behind ending a life. By retiring an android, Rick hasn't caused pain and suffering to a living thing, but rather dismissed it from its role as a servant.

It's explicitly apparent that the humans in the novel see androids as mere machines, however as previously stated, androids are practically identical to humans both physically and psychologically. They think and feel as we do, and this poses the question: if a machine functions identically to how a human, does that not then make the machine human, too?

In computer scientist Jeff Hawkins' book *On Intelligence*, the author writes that there's a fundamental difference between how computers and the brain function:

"AI suffers from a fundamental flaw in that it fails to adequately address what intelligence is or what it means to understand something." (Jeff Hawkins, 2004, p.13)

Therefore, according to Hawkins, computers as we know them will never be intelligent because they lack key features that the human brain utilises. Indeed, the human brain is incredibly complex, and Hawkins proposes that by continuing to develop AI as we are now, we will only ever create machines that *simulate* intelligence rather than genuinely possess intelligence. I think it's important to address that we ourselves also don't currently fully understand our own intelligence, nor do we know what it 'means' to understand things. Therefore, my counterargument to Hawkins' statement is that if we manage to create AI that simulates intelligence to the point of being indistinguishable from a human, then that AI could be considered human, or at the very least intelligent on some level, as well. This is the philosophy that is widely employed by developers of AI in the form of the Turing test.

Initially proposed under the name 'the imitation game' by computer scientist Alan Turing in his paper *Computing Machinery and Intelligence* (Turing, 1950), the Turing test is a method of determining the intelligence of AI systems by using a human to interact with an AI. Today, we use a more advanced method compared to Turing's initial proposal: a human participant is asked to communicate with an AI using a digital chat system, however the human doesn't know that they're talking to an AI. If the human believes that they were talking to another person over the course of the test, then the AI passes. Presently, no AI has been able to pass the test to a satisfactory degree, although in his paper Turing predicts that we will eventually arrive at a point where machines will have the capacity to fool humans into thinking they are intelligent:

"I believe that in about fifty years' time it will be possible to programme computers, with a storage capacity of about 10^9 to make them play the imitation game so well that an average interrogator will not have more than 70 percent chance of making the right identification after five minutes of questioning." (Turing, "Computing Machinery and Intelligence" *Mind*, vol. 59, p.442)

Do Androids Dream of Electric Sheep? Takes the concept of the Turing test and flips it on its head, with the human bounty hunters testing the androids instead. This is called the Voigt-Kampff test - a series of questions that are asked whilst the subject is under analysis from a machine designed to test the subject's ability to feel empathy. Verbal responses don't affect the test, but rather the body's physiological reaction to the questions. If the subject fails the test, that is, if they fail to demonstrate empathy, then they are considered androids and will be retired.

We know, in the real world, that there are people that feel no empathy. Individuals that might display psychopathic characteristics or other anti-social qualities are still human beings, making the use of empathy as a determining factor a flawed concept in the novel. In fact, as a character, Rick Deckard himself fails to feel empathy for the androids he destroys, placing him in an uncomfortable grey area with regards to his status as a human in this context.

Deckard's indifference towards the suffering of androids puts him at odds with his ownership of an electric sheep. In this post-nuclear world, all animals have become either extinct or incredibly rare, making the manufacture of robotic animals a hugely lucrative business. In order to fill the void left by naturally occurring animals, people purchase robotic ones and treat them as pets. This is a philosophical conflict; if replicants are feared and mistreated because they are artificially intelligent, ostensibly 'pretending' to be human, why then would owning an electric sheep that functions in a similar way be an acceptable replacement for a real one? This may be a product of the uncanny valley, a phenomenon where a sense of eeriness and unease is felt when observing and interacting with something that almost resembles a human, but has slight imperfections. Initially explored by robotics professor Masahiro Mori in 1970, he posits that this phenomenon may in fact be a hugely important barrier to overcome when designing machines that we as people can relate to, and suggests abandoning the concept of a human-like appearance altogether.

"We should begin to build an accurate map of the uncanny valley so that through robotics research we can begin to understand what makes us human. This map is also necessary to create—using nonhuman designs—devices to which people can relate comfortably." (Mori, MacDorman and Kageki, 2012, p.100)

Although the androids in *Do Androids Dream of Electric Sheep* are physically indistinguishable from humans, the uncanny valley could perhaps be applied to their deeper functions. Electric animals are acceptable, but once AI starts to replicate human behaviour, the people in this world respond with hostility. It's too personal, too close to home, to see a machine behave, or indeed 'pretend' to behave, as intelligently as humans do. Extending this to the real world, this may become a barrier for us, too; if we are aware that the machines we create are artificially intelligent, will we ever come to consider them as genuinely intelligent beings or forever view them as mechanical pretenders?

Bibliography

- 1 Dick, P.K., 1999. *Do Androids Dream of Electric Sheep?*, London: Gollancz.**
- 2 Hawkins, Blakeslee & Blakeslee, Sandra, 2004. *On Intelligence : [How a New Understanding of the Brain Will Lead to the Creation of Truly Intelligent Machines]*, New York: Times Books/Henry Holt.**
- 3 Turing, A.M., 1950. *Computing Machinery and Intelligence*. *Mind*, 59(236), pp.433–460.**
- 4 Mori, M., MacDorman, K. and Kageki, N., 2012. *The Uncanny Valley [From the Field]*. *IEEE Robotics & Automation Magazine*, 19(2), pp.98-100.**

